

Symcloud

Filehostingplattformen sind in der heutigen Zeit allgegenwärtig. Ohne einen Zugriff zu einem der allgemein verfügbaren Dienste, ist heutzutage eine Zusammenarbeit in einer Gruppe von Menschen zum Beispiel an einer Universität kaum möglich.

In den letzten Jahren wurde allerdings, nicht zuletzt durch den Wistle Blower Edward Snowden, bekannt dass genau diese Daten von verschiedensten Quellen abgesaugt und verarbeitet werden.

Diese Angst vor Kontrollverlust bewegt viele Menschen dazu Ihre Daten am liebsten bei sich zu behalten und mit Quelloffenen Lösungen die Daten auf Ihrem eigenen Server zu speichern.

TODO Vorstellung? und Agenda

Einführung

Meine Masterarbeit beschäftigt sich mit der Konzeption einer Speicherlösung, die Ideen aus verschiedenen Applikationen und Technologien kombiniert, die es seinen Anwendern ermöglicht Ihre Daten lokal und soher zu speichern. Diese aber bei Bedarf gezielt an andere Benutzer weiterzugeben, selbst wenn diese nicht auf dem selben Server registriert sind.

Die Idee entstand vor ein paar Jahren als ich auf ein anderen Open-Source Projekt namens ownCloud gestoßen bin und reifte dann in den letzten Monaten zu meiner Masterarbeit heran. Zuerst ging es bei dem Projekt lediglich um eine Plattform die es ermöglicht Dateien zu verwalten und mit anderen zu teilen. Bei der Recherche jedoch erkannte ich, dass die Unabhängigkeit von einem Provider sehr wichtig für die Zielgruppe des Projektes ist. Jeder der verwender will sein eigener Provider sein. Jedoch ist es dann bei klassischen Systemen wie ownCloud oder anderen nur schwer möglich Dateien über die Grenzen des eigenen Servers zu teilen.

Daher wurde der Fokus zu Beginn der Projektzeit sehr schnell der Fokus auf genau diesen Anwendungsfall gelegt.

Auf der suche nach Referenzprojekten bin ich schnell auf zwei sehr starke Inspirationsquellen gestoßen. Beide Projekte sind in ihrer Zeit gesehen Pilotprojekte und vorreiter ihrer Technologien.

Projekt Xanadu

Das Projekt Xanadu wurde durch Ted Nelson in den 1960er Jahren initiiert und ist bis heute nicht finalisiert worden. Jedoch prägte Ted Nelson mit der Gründung des Projektes und durch die Veröffentlichung des wissenschaftlichen Artikels "The Hypertext" im Jahre 1965 den Begriff des Hypertext und inspirierte damit unter anderem Tim Berners-Lee zu der entwicklung des WorldWideWeb mehr als zwei Jahrzehnte später.

Die Kernaussage des Projektes beinhalten 12 Thesen, die zum Teil im heutigen Web Anwendung finden. Jedoch meinte Ted Nelson angesprochen auf das Web in den 1990er Jahren, dass die wichtigsten Thesen eben sträflich vernachlässigt wurden.

TODO Aussage

Diaspora

Diaspora ist ein Dezentrales Soziales Netzwerk. Es ist vergleichbar mit Facebook welches jedoch zentralisiert aufgebaut ist. Was bedeutet, dass jeder Benutzer der mit anderen kommunizieren will auch bei Facebook registriert sein muss.

Diaspora geht hier einen anderen Weg. Über ein spezielles Protokoll können z.B. Nachrichten auch mit Benutzern ausgetauscht werden, die auf einem anderen Diaspora Server registriert sind.

Anforderungen

Die Anforderungen wurden in drei Teilgebiete unterteilt. Die jeweils später im Konzept eingebaut wurden.

1. Datensicherheit
2. Filehosting und Filesharing Funktionalitäten
3. Architektur

Zusätzlich wurde darauf geachtet, dass die Entwicklung auf dem aktuellen Stand der Technik und mit einem Fokus auf Wartbarkeit und Erweiterbarkeit umgesetzt wird.

Evaluierung

Für die Evaluierung bestehender Technologien, wurden moderne Anwendungen und Technologien anhand verschiedener Kriterien untersucht und geeignete Ideen für das Projekt gefunden, die dann in das Konzept miteingeflossen sind.

Konzept

Das Ergebnis der Arbeit ist ein Konzept der verteilten Datenhaltung auf Basis eines verteilten Datenmodells. Dieses Konzept ist unabhängig von der Anwendung, in der es verwendet wird und kann daher als Bibliothek in verschiedenste Anwendungen bzw. Plattformen integriert werden.

Das verteilte Datenmodell wird über ein Primärbasiertes Backup Protokoll an die Server verteilt, die Zugriff auf die Datei besitzen. Das bedeutet, dass mindestens einer der registrierten Benutzer Zugriff auf die Datei besitzen. Um die Datensicherheit zu erhöhen, können die Daten auch an vertrauenswürdige Server verteilt werden, die eigentlich keinen Zugriff auf die Daten besitzen. Die Anwendung verhindert dann den Zugriff von unbefugten Servern durch ein Rechtssystem. Dadurch können die Daten bei einem Ausfall des eigenen Servers wiederhergestellt werden.

TODO Ergebnis aussagen?

Implementierung

Bei der Implementierung wurde auf bewährte Technologien aus dem Umfeld von PHP und dem Framework Symfony2 vertraut. Auf diese Technologien möchte ich hier nicht genauer eingehen. Dies würde den Rahmen dieser Präsentation sprengen.

Entstanden ist während der Projektzeit ein funktionstüchtiger Prototyp, der die funktionalen Anforderungen erfüllt.

Es wurde eine Bibliothek implementiert, die eine einfache verteilte Datenbank enthält, die die Daten lokal in einem Ordner ablegt und über das Verteilungsprotokoll die Daten an andere bekannte Server repliziert. Diese Daten können dann von dem anderen Server verwendet werden.

Diese Bibliothek wurde in eine bestehende Content Management Plattform integriert. Über eine einfache Benutzerschnittstelle, können die Benutzer die Daten bearbeiten bzw. erstellen.

Als zusätzliche Komponente wurde ein Synchronisierungstool umgesetzt, mit dem es möglich ist die Daten von einem lokalen Ordner mit einem Server zu synchronisieren. Dieses Tool erkennt automatisch welche Dateien aktualisiert, erstellt oder gelöscht werden müssen. Diese Funktionalitäten funktionieren bidirektional. Das bedeutet, dass das Tool beidseitig Änderungen erkennen kann.